

# An Analysis of Low Complexity Algorithms for MIMO Antenna Selection

Peter J. Smith<sup>†</sup>

Tim W. King<sup>†</sup>

Lee M. Garth<sup>†</sup>

Mischa Dohler<sup>\*</sup>

<sup>†</sup> Department of Electrical and Computer Engineering, University of Canterbury, Christchurch, New Zealand

Email: {p.smith,twk20,l.garth}@elec.canterbury.ac.nz

<sup>\*</sup> France Télécom R&D, 28 Chemin du Vieux Chêne, 38243 Meylan Cedex, France

Email: mischa.dohler@francetelecom.com

**Abstract**—In this paper we consider transmit and receive selection methods designed to achieve high channel capacities in a single-user MIMO link. A variety of radio channels are considered, including i.i.d. Rayleigh, correlated Rayleigh and Ricean fading environments. Also considered is the presence of imperfect channel state information (CSI) and a simplified waterfilling scheme. In all cases, we evaluate the performance of optimal selection, simple norm based selection and other benchmark selection techniques. The major contribution is a general approach to analyzing the capacity of the selection schemes via a simple power scaling factor. We are able to assess the impact of different channels, imperfect CSI and power allocation using this power scaling factor. Furthermore, the analysis is valid for all scenarios: transmit selection, receive selection and joint transmit-receive selection. Results are shown which compare the capacity performance over a wide range of cases. A notable conclusion is that optimal selection, which is computationally intensive, is outperformed at low SNR by the simple, norm based approach with power allocation.

## I. INTRODUCTION

Given the advantages of wireless Multiple-Input Multiple-Output (MIMO) systems, using transmit (TX) and receive (RX) antenna arrays, much attention has been paid to these systems and their potential limits. As the system sizes are increased to utilize the potential of MIMO techniques, the increased hardware requirements become a key issue. Multiple RF chains are needed to optimize performance at the expense of cost and volume. In this paper we consider a simple yet powerful approach for reducing the system complexity whilst retaining a considerable portion of the increased rates from using MIMO technology. Using the principle of antenna selection at the TX and/or RX side, RF chains can be optimally assigned to a subset of the available antennas. This provides greater spatial multiplexing gain without a large increase in the system's hardware needs.

Many researchers have proposed algorithms for antenna selection [1]–[3]. Some of these approaches focus on selection to minimize outage probabilities or error rates [4], [5]. We consider selection methods which target high capacity values as in [2], [6]. Previous work in this area includes simple methods [1], [3], [7], variants of which are discussed in this paper, as well as complex methods [8], which may give slightly better performance than the simple methods but with much greater computational complexity. In this paper we propose to build on the ideas in [1]–[3], analyzing and extending the algorithms

that were discussed. We demonstrate in our statistical analysis of the resulting channel capacities that the selection process can be modeled as a simple scaling of the SNR, resulting in a performance gain. Thus, we propose the SNR scaling factor as a very simple metric which can be used to compare the performances of the algorithms. Note that this approach is extremely general and can be used for a variety of channels, selection scenarios, and also when channel estimation errors are considered.

This paper is organized as follows. In Sec. II we review the MIMO channel model, the channel capacity for antenna selection systems and a model for imperfect channel state information (CSI). We introduce our various selection algorithms in Sec. III, followed by a statistical analysis of the algorithms in Sec. IV. Then we explore the effects of imperfect channel knowledge in Sec. V. We verify our analysis using Monte Carlo simulations in Sec. VI. We conclude in Sec. VII.

## II. MIMO CHANNEL MODEL AND CAPACITY

We focus on a single user MIMO link with  $n$  transmitters and  $m$  receivers with channel observations of the form

$$\mathbf{r} = \mathbf{H}\mathbf{u} + \mathbf{n} \quad (1)$$

where vectors  $\mathbf{r}$ ,  $\mathbf{u}$ , and  $\mathbf{n}$  contain the received observations, the transmitted symbols and the channel noise, respectively. Without a loss of generality we assume that  $E\{|n_i|^2\} = 1$ . The  $m \times n$  channel matrix  $\mathbf{H}$  is used to model a flat-fading Gaussian channel with elements  $h_{ij}$  having complex Gaussian distributions. We focus on the i.i.d. Rayleigh case but do discuss semi-correlated (SC) Rayleigh and Ricean channels. We assume that this CSI, whether perfect or imperfect, is known at both the transmitter and receiver. The signal-to-noise ratio (SNR) has the form  $E\{\|\mathbf{u}\|^2\}$ . If  $q = \min(m, n)$ , the channel capacity for this system is then

$$C_{\text{full}} = \log_2 \left| \mathbf{I}_q + \frac{\text{SNR}}{n} \mathbf{W} \right| \quad (2)$$

where  $|\cdot|$  represents the matrix determinant,  $\mathbf{I}_q$  is the  $q \times q$  identity matrix, and the  $q \times q$  matrix  $\mathbf{W}$  is given by

$$\mathbf{W} = \begin{cases} \mathbf{H}\mathbf{H}^\dagger, & \text{for } m \leq n \\ \mathbf{H}^\dagger\mathbf{H}, & \text{for } m > n \end{cases} \quad (3)$$

where  $(\cdot)^\dagger$  denotes the conjugate transpose. To simplify the notation and without loss of generality, from now on we assume that the number of columns is not less than the number of rows in the channel or selection matrices we define.

If we choose to use only  $r$  of the receive antennas and  $t$  of the transmit antennas, we are left with a  $r \times t$  submatrix  $\mathbf{S}$  of the main channel matrix  $\mathbf{H}$ . The channel capacity then has the form

$$C_{\text{sel}} = \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{S} \mathbf{S}^\dagger \right|. \quad (4)$$

We now define the unordered column norms of  $\mathbf{H}$  to be  $P_c^{(u)} = \sum_{i=1}^m |h_{ic}|^2$ , for  $c = 1, \dots, n$ , and the unordered row norms to be  $Q_r^{(u)} = \sum_{j=1}^n |h_{rj}|^2$ , for  $r = 1, \dots, m$ . Ordering the column norms of  $\mathbf{H}$  yields the ordered set  $\{P_{1:n} > P_{2:n} > \dots > P_{n:n}\}$ , where  $P_{j:n}$  denotes the  $j$ -th largest column norm. We also define  $\mu_{j:n}$  to be the mean of  $P_{j:n}$ . For ease of notation in the rest of the paper we use  $P_j$  in place of  $P_{j:n}$  and  $\mu_j$  in place of  $\mu_{j:n}$ .

In the case of imperfect CSI the estimated channel matrix  $\widehat{\mathbf{H}}$  is commonly modelled as

$$\mathbf{H} = \rho \widehat{\mathbf{H}} + \sqrt{1 - \rho^2} \mathbf{E} \quad (5)$$

where correlation  $\rho = \text{corr}(h_{ij}, \widehat{h}_{ij})$  and  $\mathbf{E}$  is an error matrix which is statistically identical to  $\mathbf{H}$ . The column and row norms for  $\widehat{\mathbf{H}}$  are defined as before with  $\widehat{P}_c^{(u)} = \sum_{i=1}^m |\widehat{h}_{ic}|^2$ , for  $c = 1, \dots, n$ , and  $\widehat{Q}_r^{(u)} = \sum_{j=1}^n |\widehat{h}_{rj}|^2$ , for  $r = 1, \dots, m$ .

### III. SELECTION ALGORITHMS

We now describe a variety of existing and novel transmit and/or receive selection algorithms. All of these algorithms involve selecting a  $r \times t$  submatrix  $\mathbf{S}$  from the  $m \times n$  channel matrix  $\mathbf{H}$ . Equations (2) and (4) define the capacity relationships for the full and selection channels, respectively. We focus on the scenarios where selection is performed either at the TX or at both the TX and RX. Analogous RX-only selection methods can easily be obtained from the respective TX selection methods.

#### A. Algorithms for Transmit Selection

The optimal selection method involves computing the channel capacity for all possible submatrices of  $\mathbf{H}$  and picking the  $\mathbf{S}_{\text{OSA}}$  which yields the highest channel capacity. We can write this Optimal Selection Algorithm (OSA) as

$$\mathbf{S}_{\text{OSA}} = \arg \max_{\{\mathbf{S} \in \mathbf{H}\}} \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{S} \mathbf{S}^\dagger \right| \quad (6)$$

where set  $\{\mathbf{S} \in \mathbf{H}\}$  contains all  $m \times t$  submatrices of  $\mathbf{H}$ .

The second case is named the Arbitrary Selection Algorithm (ASA) in which an arbitrary  $\mathbf{S}_{\text{ASA}}$  is chosen. The simplest way to do this is to take the first  $t$  out of the  $n$  columns of  $\mathbf{H}$ .

The OSA requires an exhaustive search over all possible submatrices of  $\mathbf{H}$ , yielding a large computational complexity. As a simple suboptimal alternative, consider forming  $\mathbf{S}_{\text{NSA}}$  using the columns of  $\mathbf{H}$  with the  $t$  biggest column norms,

$P_j$ ,  $j = 1 \dots t$ . Many other researchers [1], [3] have proposed this simple selection method, which we call the Norm-Based Selection Algorithm (NSA). We will show, using analysis and simulations, that the capacity for the NSA comes close to that for the OSA.

#### B. Algorithms for Transmit-Receive Selection

All the TX selection algorithms (OSA, ASA and NSA) can be extended for RX antenna selection. The main difference is that the calculations are now over the rows of  $\mathbf{H}$  rather than the columns as in the TX selection case. When jointly implementing both TX and RX selection, the OSA method involves searches over both the rows and columns. The ASA method is equivalent to selecting the principal  $r \times t$  submatrix of  $\mathbf{H}$ . For the NSA approach, RX selection can be performed independently of TX selection (i.e., the selection of columns does not affect the selection of rows) or jointly with TX selection. In this joint case the selection of the RX antennas is performed on  $\mathbf{S}$  before or after the TX selection. Intuitively, we should first select over the dimension with the greatest choice (i.e., select TX antennas first if  $m \leq n$  and RX antennas first if  $m > n$ ). In principle we could also select the rows and columns iteratively. However, the differences observed between such methods are very small. We will show in the analysis and simulation section that joint selection using the NSA slightly outperforms independent selection with the same algorithm.

#### C. "Poor Man's Waterfilling"

Note that (2) and (4) assume equal transmit power levels among the TX antennas. For the case of TX selection, if we select  $t$  antennas with corresponding norms  $P_1, P_2, \dots, P_t$ , consider the novel power allocation method where the power is allocated to each antenna proportional to its column norm. Hence, if an antenna has column norm  $P_j$ , it is allocated a proportion  $P_j / \sum_{k=1}^t P_k$  of the total power. Hence, the norm of a vector containing the column norms becomes  $t P_j^2 / \sum_{k=1}^t P_k$ . Due to its low complexity compared to standard waterfilling techniques, we have dubbed this "Poor Man's Waterfilling" (PMWF). In Sec. IV-F we will analyze PMWF and show that in conjunction with NSA, it outperforms OSA at low SNRs.

### IV. ALGORITHM ANALYSIS

In this paper, we do not analyze the OSA method because it is an extremely difficult nonlinear problem. The analytical properties of the ASA method are well known, as the method leads to the standard capacity results [9]. Here we concentrate our analysis primarily on the NSA method. The aim of this analysis is to develop effective methods to calculate the gains made by norm-based antenna selection. Our statistical analysis is based mainly on the fact that capacity, although exactly defined by the joint distributions of the elements *all* possible submatrices of  $\mathbf{H}$ , is strongly affected by the moments of these elements. Finally, note that in Secs. IV-A–IV-D we only consider TX selection, whereas in Sec. IV-E we extend our analysis to joint TX-RX selection.

### A. Norm-Selection Algorithm

Consider the simplest case of a (2,2) Rayleigh channel matrix with column norms  $P_1^{(u)}, P_2^{(u)}$ . Assuming the first column  $[h_{11}, h_{21}]^T$  has the largest norm and we are selecting one TX antenna, then  $\mathbf{S} = [h_{11}, h_{21}]^T$ . Note that the elements of  $\mathbf{S}$  are no longer Gaussian since their distribution is conditioned on  $P_1^{(u)} > P_2^{(u)}$ . The exact distribution of  $\mathbf{S}$  can be obtained by computing the conditional joint density  $f(h_{11}, h_{21} | P_1^{(u)} > P_2^{(u)})$ . Clearly, the elements of  $\mathbf{S}$  will have a larger variance than those of  $\mathbf{H}$ , since they were selected to have the largest column norm. It is less clear how the conditioning affects the type of distribution. Some analysis for the (2,2) case and simulations for larger systems show that the elements of  $\mathbf{S}$  are statistically very similar to the original  $\mathbf{H}$  elements but with an increased variance. Certainly they remain zero mean and symmetric. The peak of the density is flattened slightly compared to the Gaussian, which is caused by the selection of large column norms, pushing more probability away from zero.

Hence, we heuristically model the effect of TX selection as a simple scaling of the original elements of  $\mathbf{H}$ . Column  $j$  of  $\mathbf{S}$  has norm  $P_j$  (after rearranging columns). Following the above approach, we model  $\mathbf{S}$  as having elements with the same distribution as those of  $\mathbf{H}$ , but with variances  $E\{P_j\}/m = \mu_j/m$  rather than 1 for column  $j$ . This gives us a power scaling due to the selection process. Hence, we approximate  $\mathbf{S}$  by

$$\mathbf{S} \approx \mathbf{V} \text{diag}(\sqrt{\mu_1}, \sqrt{\mu_2}, \dots, \sqrt{\mu_t}) / \sqrt{m} \quad (7)$$

where  $m \times t$  matrix  $\mathbf{V}$  is statistically identical to an  $m \times t$  submatrix of  $\mathbf{H}$  picked at random. Defining  $\mathbf{M} = \text{diag}(\mu_1, \mu_2, \dots, \mu_t)/m$ , equation (7) leads to the approximate capacity

$$C_{\text{sel,pa}} \approx \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{V} \mathbf{M} \mathbf{V}^\dagger \right|. \quad (8)$$

where the subscript, “pa”, denotes the fact that the diagonal matrix,  $\mathbf{M}$ , can be interpreted as performing power allocation over the antennas. We refer to this technique as the power allocation (PA) approximation.

We can further approximate (8) by replacing the power allocation matrix  $\mathbf{M}$  by a single power scaling factor

$$P_{\text{av}} = \frac{1}{m t} \sum_{j=1}^t \mu_j. \quad (9)$$

This leads to the capacity

$$C_{\text{sel,ps}} \approx \log_2 \left| \mathbf{I}_r + P_{\text{av}} \frac{\text{SNR}}{t} \mathbf{V} \mathbf{V}^\dagger \right| \quad (10)$$

and a very simple interpretation of the NSA as providing a power scaling of  $P_{\text{av}}$ . This technique is referred to as a power scaling (PS) approximation. To implement both the PA and PS approximations, we require  $\mu_1, \mu_2, \dots, \mu_t$  and the statistics of (8) and (10) for the various channels. Note that the mean capacity of (10) is known for all the channels considered here [9]–[11]. We now consider these statistics for a variety of channel distributions.

### B. Rayleigh Flat-Fading Channels

The Rayleigh channel defined in Sec. II has Gaussian matrix elements  $h_{ij}$ , and thus its column norm powers  $P_j^{(u)}$  are i.i.d. complex  $\chi^2$  distributed with  $m$  degrees of freedom<sup>1</sup> and a mean value of  $E\{P_j^{(u)}\} = m$ . The complex  $\chi^2$  probability density function (pdf) and the cumulative distribution function (cdf) are, respectively

$$f_\chi(x; m) = \frac{x^{m-1}}{(m-1)!} \exp(-x) \quad (11)$$

$$F_\chi(x; m) = 1 - \exp(-x) \sum_{k=0}^{m-1} \frac{x^k}{k!}. \quad (12)$$

From (11) and (12) we can calculate the means of the ordered column norms using simple order statistics. Hence, we can write [12]

$$\mu_j = \binom{n}{j} \int_{-\infty}^{\infty} x F(x)^{j-1} [1 - F(x)]^{n-j} f(x) dx. \quad (13)$$

The combination of (13) with (11) and (12) gives a complicated integral expression which can be solved analytically but is cumbersome. We prefer to evaluate (13) numerically since it converges quickly using a numerical integration routine. In future work, we plan to investigate the use of quantile approximations to  $\mu_j$  [12] which lead to even simpler, closed-form expressions.

### C. Semi-Correlated Rayleigh Flat Fading Channels

Consider now a SC Rayleigh channel matrix given by:

$$\mathbf{H}_{\text{sc}} = \mathbf{R}^{1/2} \mathbf{H} \quad (14)$$

where  $\mathbf{R}^{1/2}$  represents the spatial correlation at the the receiver. The column norms of  $\mathbf{H}_{\text{sc}}$  are now given by  $P_j^{(u)} = \mathbf{h}_j^\dagger \mathbf{R} \mathbf{h}_j$ ,  $j = 1, \dots, n$ , where  $\mathbf{h}_j$  is the  $j$ -th column of  $\mathbf{H}$ . This is a well known quadratic form [13] and can be rewritten as  $P_j^{(u)} = \sum_{i=1}^m \lambda_i |h_{ij}|^2$  where  $\lambda_1, \dots, \lambda_m$  are the eigenvalues of  $\mathbf{R}$ . The density and distribution of  $P_j^{(u)}$  are, respectively

$$f(x) = \sum_{j=1}^m \frac{b_j}{\lambda_j} \exp\left(-\frac{x}{\lambda_j}\right) \quad (15)$$

$$F(x) = \sum_{j=1}^m b_j \left[ 1 - \exp\left(-\frac{x}{\lambda_j}\right) \right] \quad (16)$$

where

$$b_j = \lambda_j^{m-1} \sum_{k=1, k \neq j}^m (\lambda_j - \lambda_k)^{-1}.$$

As before, the capacity approximations (8) and (10) can be extended to this case. Note that  $\mathbf{V}$  remains semi-correlated. If

<sup>1</sup>Note that this is slightly different from a real  $\chi^2$  distribution with  $2m$  degrees of freedom.

we define  $\mathbf{V} = \mathbf{R}^{1/2}\mathbf{U}$ , where  $\mathbf{U}$  is an i.i.d. Rayleigh channel matrix, then the capacity approximations can be rewritten as

$$C_{\text{sel,pa}} \approx \log_2 \left| \mathbf{I}_r + \frac{\text{SNR}}{t} \mathbf{R}^{1/2} \mathbf{U} \mathbf{M} \mathbf{U}^\dagger \mathbf{R}^{1/2} \right| \quad (17)$$

$$C_{\text{sel,ps}} \approx \log_2 \left| \mathbf{I}_r + P_{\text{av}} \frac{\text{SNR}}{t} \mathbf{R}^{1/2} \mathbf{U} \mathbf{U}^\dagger \mathbf{R}^{1/2} \right|. \quad (18)$$

Again, we can derive  $\mu_1, \dots, \mu_t$  and, hence,  $\mathbf{M}$  and  $P_{\text{av}}$  using order statistics, as (13) and (9) are perfectly applicable to this situation, with the  $\chi^2$  distribution replaced by the quadratic form distribution in (16). Again, analytical results for the  $\mu_j$ 's are possible, but we prefer to evaluate (13) numerically.

#### D. Ricean Channels

An i.i.d. Ricean channel can be expressed as the sum of a line-of-sight (LOS) channel and a Rayleigh flat-fading channel as follows

$$\mathbf{H}_{\text{rice}} = a \mathbf{H}_{\text{LOS}} + b \mathbf{H} \quad (19)$$

where  $a^2 + b^2 = 1$  and the entries of  $\mathbf{H}_{\text{LOS}}$  and  $\mathbf{H}$  have unit variance. If we assume the common model of a unit rank matrix  $\mathbf{H}_{\text{LOS}}$ , then we can assume w.l.o.g. that the elements of  $\mathbf{H}_{\text{LOS}}$  are all unity. Hence the  $j$ -th column norm of  $\mathbf{H}_{\text{rice}}$  is given by  $P_j^{(u)} = [a \mathbf{1} + b \mathbf{h}_j]^\dagger [a \mathbf{1} + b \mathbf{h}_j]$  where  $\mathbf{1} = [1, \dots, 1]^T$  and  $\mathbf{h}_j$  is again the  $j$ -th column of the i.i.d. Rayleigh matrix  $\mathbf{H}$ . This is a non-central quadratic form, and its pdf and cdf are respectively given by [13]

$$f(x) = \frac{1}{b^2} \left( \frac{2x}{b^2\delta} \right)^{(m-1)/2} I_{m-1} \left( \frac{2x\sqrt{\delta}}{b^2} \right) \times \exp \left[ - \left( \frac{2x}{b^2} + \delta \right) / 2 \right] \quad (20)$$

$$F(x) = \exp \left( -\frac{\delta}{2} \right) \sum_{j=0}^{\infty} \frac{(\delta/2)^j}{j! (m+j-1)! 2^{m+j}} \times \int_0^{2x/b^2} u^{m+j-1} \exp \left( -\frac{u}{2} \right) du \quad (21)$$

where  $\delta = 2m a^2/b^2$  and  $I_k(\cdot)$  is a modified Bessel function. Note that the integral in the cdf can be given as a finite sum, but there is no closed-form solution for the cdf which avoids either numerical integration or an infinite series. Again, (13) can be used to obtain  $\mu_1, \mu_2, \dots, \mu_t$  and the corresponding capacity approximations (7) and (10) can be computed.

#### E. Transmit-Receive Antenna Selection

As mentioned in Sec. III-B, the TX selection algorithms can be extended for RX antenna selection. We denote the TX and RX selection gains by  $P_{\text{av},t}$  and  $P_{\text{av},r}$ , respectively. If we perform joint TX and RX selection, we can leverage the results from Secs. IV-A and IV-D to approximate the capacity. Assuming that TX selection occurs first, we have the selected matrix  $\mathbf{S}_t \approx \mathbf{V} \mathbf{M}^{1/2}$  as in (7). For the i.i.d. Rayleigh case, the row selection process is then exactly the same as TX selection for the SC Rayleigh channel, since the form of the rows of  $\mathbf{S}_t$  match the columns of  $\mathbf{H}_{\text{sc}}$  in (14). Hence, we can evaluate both (8) and (10) for the i.i.d. Rayleigh case.

For the SC Rayleigh and i.i.d. Ricean channels similar methods also lead to capacity approximations of the form given in (8). For space reasons these derivations are omitted here. However, the SNR scaling approximation (10) is simple to derive for all channels. After TX selection we have  $\mathbf{S}_t \approx P_{\text{av},t}^{1/2} \mathbf{V}$ . Similarly, after RX selection we have  $\mathbf{S} \approx P_{\text{av},r}^{1/2} \mathbf{S}_t = P_{\text{av},r}^{1/2} P_{\text{av},t}^{1/2} \mathbf{V}$ . Hence, we can use (10) with the new power scaling  $P_{\text{av}} = P_{\text{av},r} P_{\text{av},t}$ .

#### F. Poor Man's Waterfilling (for TX Selection)

After the initial antenna selection, other techniques can be used to further enhance the capacity of the system. These include techniques like conventional waterfilling, where the transmit power is allocated along the eigenmodes of the channel. We propose a new simpler algorithm to achieve suboptimal gains, but at a highly reduced complexity level.

"Poor Man's Waterfilling" (PMWF) employs the values of the column norms  $P_j$  to allocate power to the system. We focus on a power scaling approximation to the effect of PMWF. After TX selection has been performed, column  $j$  in  $\mathbf{S}$  has norm  $P_j$ . Using PMWF this norm is scaled by  $tP_j / \sum_{k=1}^t P_k$  and the resulting matrix is denoted  $\mathbf{S}_{\text{pmwf}}$ . The total norm of  $\mathbf{S}_{\text{pmwf}}$  is  $\|\mathbf{S}_{\text{pmwf}}\|^2 = \sum_{j=1}^t tP_j^2 / \sum_{k=1}^t P_k$ . The average modulus squared value of any element in  $\mathbf{S}_{\text{pmwf}}$  is therefore  $\|\mathbf{S}_{\text{pmwf}}\|^2 / (mt)$ , and the mean value of this gives the power scaling factor:

$$P_{\text{av,pmwf}} = \frac{1}{m} \mathbb{E} \left\{ \frac{\sum_{j=1}^t P_j^2}{\sum_{k=1}^t P_k} \right\}. \quad (22)$$

For TX-RX selection using PMWF we have  $P_{\text{av}} = P_{\text{av},r} P_{\text{av,pmwf}}$ . At present (22) has to be evaluated via simulation. An analytical investigation of (22) is part of ongoing work.

#### V. EFFECTS OF IMPERFECT CHANNEL STATE INFORMATION

In this section we assess the impact of imperfect CSI on TX selection via a simple power scaling approach. Taking column  $j$  of (5) gives

$$\mathbf{h}_j = \rho \hat{\mathbf{h}}_j + \sqrt{1 - \rho^2} \mathbf{e}_j \quad (23)$$

with  $\mathbf{h}_j$ ,  $\hat{\mathbf{h}}_j$ ,  $\mathbf{e}_j$  being the  $j^{\text{th}}$  columns of  $\mathbf{H}$ ,  $\hat{\mathbf{H}}$  and  $\mathbf{E}$  respectively. Taking the norms of the columns in (23) gives

$$P_j^{(u)} = \rho^2 \hat{P}_j^{(u)} + (1 - \rho^2) \|\mathbf{e}_j\|^2 + \Delta \quad (24)$$

where  $\Delta$  is the cross product term. Since  $\mathbb{E}\{\hat{P}_j^{(u)}\} = \mathbb{E}\{\|\mathbf{e}_j\|^2\} = m$ , we can rewrite (24) as

$$P_j^{(u)} = \rho^2 (\hat{P}_j^{(u)} - m) + m + \epsilon_j \quad (25)$$

where  $\epsilon_j = \Delta + (1 - \rho^2)(\|\mathbf{e}_j\|^2 - m)$ . In the process of TX selection we are selecting columns of  $\mathbf{H}$  on the basis of measurements of  $\hat{\mathbf{H}}$  which leads to the area of order statistics called concomitants [12]. Note that (25) is a simple linear

model, and in the situation where  $\epsilon_j$  is independent of  $\hat{P}_j^{(u)}$  we have [12]

$$\mu_{[j]} = \rho^2 (\mu_j - m) + m \quad (26)$$

where  $\mu_j$  is as before but  $\mu_{[j]}$  represents the mean norm of the column of  $\mathbf{H}$  selected on the basis of  $\hat{\mathbf{H}}$ . In our scenario,  $\epsilon_j$  is uncorrelated with  $\hat{P}_j^{(u)}$  but not independent. Hence we apply (26) as an approximate result. Since the column norms now have means  $\mu_{[j]}$  we can adjust  $P_{av,t}$  to give  $P_{av,CSI} \triangleq \sum_{j=1}^t \mu_{[j]} / (mt)$  which can be rewritten as

$$P_{av,CSI} = 1 - \rho^2 + \rho^2 P_{av,t}. \quad (27)$$

Again the effect of another system factor, i.e. CSI, is simply accommodated in the power scaling factor.

## VI. RESULTS

Figure 1 shows a comparison of the key selection algorithms OSA, NSA, ASA for three different channels: i.i.d. Rayleigh, SC Rayleigh and Ricean. All curves are for a SNR of 10dB and show capacity vs.  $N$  in a  $(N, N+2)$  system where an  $(N, N)$  subsystem is to be selected. Parameter values are selected for ease of comparison rather than physical reasons. The SC Rayleigh channel has an exponential correlation structure [14] at the receiver with  $r = 0.8$  giving the correlation between adjacent antennas. The Ricean channel has K-factor,  $K = 10$ dB and is defined in [11]. Figure 1 shows that NSA offers a reasonable proportion of the benefits of OSA at a much reduced complexity. This holds true for all of the channel models.

Figure 2 shows the PA and PS approximations to NSA for the same three channels and SNR value as in Fig. 1. Note that both approximations are very accurate and PS performs at least as well as PA for all cases. Hence, since PS is simpler we concentrate on this approximation in the following simulations.

Results in Figs. 1 and 2 show performance at a SNR of 10dB. In Fig. 3 we look at the low SNR case (SNR = 0dB) and show that NSA with PMWF is superior to OSA here. In addition, the PS approximation is excellent and so we may also approximate the OSA performance using our simple power scaling approach for the mixture of NSA and PMWF. We simulate an i.i.d. Rayleigh channel. Note that NSA with PMWF outperforms OSA for low SNR, but for SNRs between 5 and 10dB OSA becomes superior again, although the difference is small.

In (27) we showed a quadratic relationship between the PS factor and  $\rho$ , which is a measure of the CSI quality. This result is verified in Fig. 4 which shows the PS approximation lower bounding the simulated values and a quadratic shape. Results here are for two scenarios and the plot is of the relative capacity increase due to NSA to show both cases on the same scale. Note the increased benefits due to selection as the number of redundant antennas increases from 2 to 4. Again the low SNR region is considered.

Finally, in Fig. 5 we look at the behavior of joint TX-RX selection at a SNR of 10dB. As shown, the benefit of joint

selection at the TX and RX rather than independent selection at both ends is very slight. Again, the PS approximation is excellent and the improvement over TX selection is only marginal.

The NSA and PMWF by design are low complexity algorithms involving the computation of column norms and ordering only. OSA, on the other hand, is highly intensive requiring  $\binom{m}{r} \times \binom{n}{t}$  matrix determinants, in addition to ordering and other matrix operations. Hence, the superior performance of OSA with PMWF at low SNR is quite notable.

## VII. CONCLUSION

The NSA is seen to be an important selection method, since its complexity is far less than OSA, its performance is close to OSA and, with the addition of PMWF, it can outperform OSA at low SNR levels. Hence, an analysis method for NSA is important and the power scaling method we have developed is able to approximate capacity for NSA over a variety of channels and system scenarios. The beauty of the PS approach is its generality and its ability to gauge the effect of selection in a single number. In particular, analytical results clearly show the effects of PMWF and imperfect CSI in simple closed-form expressions.

## REFERENCES

- [1] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.
- [2] R. S. Blum and J. H. Winters, "On optimum MIMO with antenna selection," *IEEE Commun. Lett.*, vol. 6, no. 8, pp. 322–324, Aug. 2002.
- [3] Z. Zhou, Y. Dong, X. Zhang, W. Wang, and Y. Zhang, "A novel antenna selection scheme in MIMO systems," in *Proc. Int'l. Conf. on Communications, Circuits and Systems*, Chengdu, China, June 27–29, 2004, pp. 190–194.
- [4] R. Heath and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. IEEE Int'l. Conf. on Communications*, Helsinki, Finland, 2001, pp. 2276–2280.
- [5] Z. Chen, B. Vucetic, and J. Yuan, "Space-time trellis codes with transmit antenna selection," *IEEE Electronics Letters*, vol. 39, pp. 854–855, May 2003.
- [6] A. Molisch, M. Win, and J. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. IEEE Int'l. Conf. on Communications*, Helsinki, Finland, 2001, pp. 570–574.
- [7] M. Gharavi-Alkhansari and A. B. Gershman, "Fast antenna subset selection in MIMO systems," *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 339–347, Feb. 2004.
- [8] A. Molisch and X. Zhang, "FFT-based hybrid antenna selection schemes for spatially correlated MIMO channels," *IEEE Commun. Lett.*, vol. 8, no. 1, pp. 36–38, Jan. 2004.
- [9] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. on Telecomm. Related Technol.*, vol. 10, pp. 585–595, Nov.–Dec. 1999.
- [10] P. J. Smith, S. Roy, and M. Shafi, "Capacity of MIMO systems with semicorrelated flat fading," *IEEE Trans. Inform. Theory*, vol. 49, no. 10, pp. 2781–2788, Oct. 2003.
- [11] M. Kang and M.-S. Alouini, "On the capacity of MIMO Rician channels," in *Proc. 40th Annu. Allerton Conf. Communications, Control, and Computing*, Allerton Monticello, IL, USA, Oct. 2002, pp. 936–945.
- [12] H. David and H. Nagaraja, Eds., *Order Statistics*. Hoboken, New Jersey, USA: John Wiley and Sons Inc., 2003.
- [13] N. L. Johnson and S. Kotz, Eds., *Continuous Univariate Distributions-2*. Hoboken, New Jersey, USA: John Wiley and Sons Inc., 1970.
- [14] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, Sept. 2001.

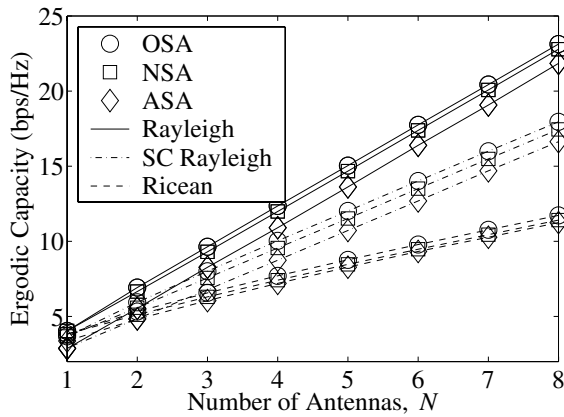


Fig. 1. Comparison of selection schemes for a  $(N, N+2)$  choose  $(N, N)$  system (SNR = 10dB).

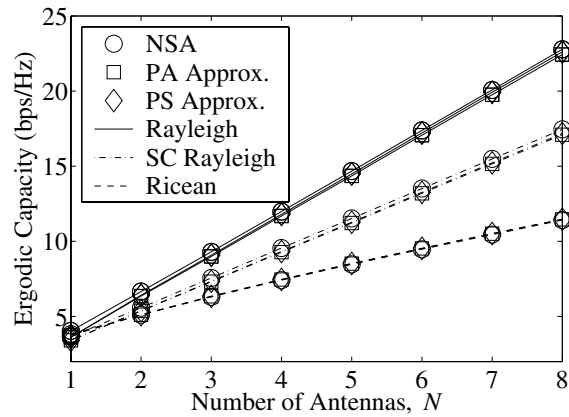


Fig. 2. Approximations for NSA for a  $(N, N+2)$  choose  $(N, N)$  system (SNR = 10dB).

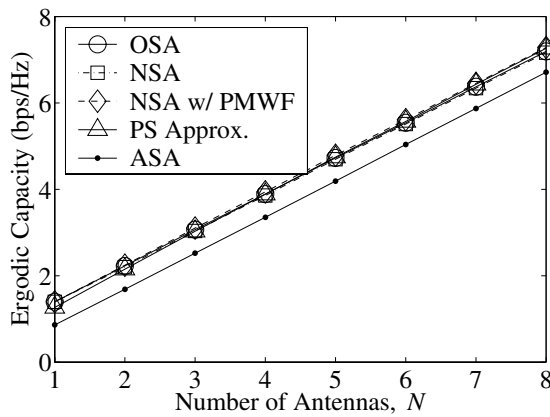


Fig. 3. PMWF comparison for a  $(N, N+2)$  choose  $(N, N)$  system for an i.i.d. Rayleigh channel (SNR = 0dB).

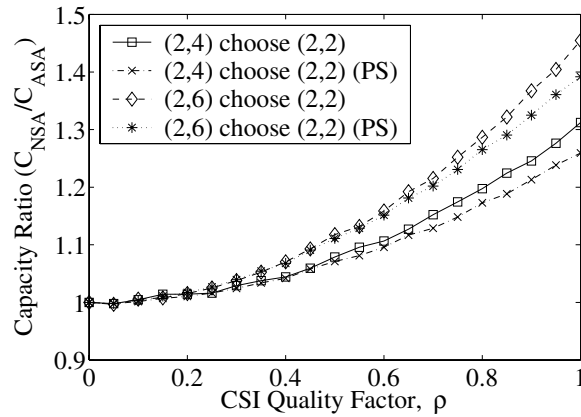


Fig. 4. Effect of imperfect channel state information on antenna selection for an i.i.d. Rayleigh channel (SNR = 0dB).

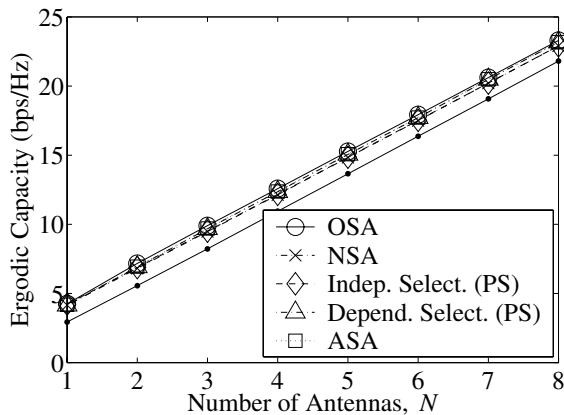


Fig. 5. TX-RX selection for a  $(N+1, N+1)$  choose  $(N, N)$  system for an i.i.d. Rayleigh channel (SNR = 10dB).